Wavelets and Filter Banks: Relationships and New Results

Martin Vetterli and Cormac Herley
Department of Electrical Engineering
and Center for Telecommunications Research
Columbia University, New York, NY 10027-6699

Abstract

The wavelet transform has recently emerged as a powerful tool for non-stationary signal analysis. Its discrete version is closely related to filter banks which have been studied in digital signal processing. Also, multiresolution signal analysis has been used in image processing. This paper indicates the relationship between these techniques. Then, it is shown how to construct biorthogonal systems with linear phase FIR filters, and having regular analysis and synthesis. Some examples of practical interest are given. The complexity of the discrete wavelet transform is also discussed.

1 Introduction

The search for basis functions better suited for the analysis of non-stationary signals is a topic of on-going research in signal processing. The short-time Fourier transform (STFT) uses basis functions of the type:

$$c_{mn}(t) = e^{-jm\omega_0 t} w(t - nt_0), \qquad (m, n \in \mathcal{Z}) \qquad (1)$$

where $w(.)$ is an appropriate window function, like a Gaussian as in the Gabor basis. Then, the STFT is the inner product of the signal with the set of basis functions. It is clear from (1) that the STFT has the same frequency and time resolution at all points (m, n) of the time/frequency plane. This can be unsatisfactory in signal analysis, where signal discontinuities should be resolved sharply in time at high frequencies while slow variations should be seen sharply at low frequencies. Of course, the time/frequency product is lower bounded by the uncertainty principle, but instead of a constant time and frequency resolution, one can trade-off one for the other. This is achieved with wavelets, where the family of basis functions is obtained by translation and dilation/contraction of a single prototype wavelet $h(t)$ [2]:

$$h_{mn}(t) = a_0^{-m/2} \cdot h(\frac{t}{a_0^m} - nb_0), \;\; (m, n \in \mathcal{Z}), \; a_0 > 1, \; b_0 \neq 0 \qquad (2)$$

where typically $a_0 = 2$ and $b_0 = 1$. The scale factor $a_0^{-m/2}$ is used to conserve the L_2 norm of the wavelet. For large positive m, $h_{mn}(t)$ is dilated by a_0^m and shifted by large steps $a_0^m b_0$, while for negative m, $h_{mn}(t)$ is contracted and shifted in small steps. Therefore, the wavelet transform is sharp in time at high frequencies as well as sharp in frequency at low frequencies (with corresponding loss of frequency or time resolution). Notice the "constant shape" property of the wavelets (due to dilation) which does not hold for the STFT (which uses modulation).

While the continuous time wavelet transform has a number of attractive features, we will be concerned with the discrete case in what follows. Such a discrete wavelet transform (DWT) is shown in figure 1 and is implemented with multirate filter banks [6]. At each stage, the spectrum is divided into low and high half band, leading to a logarithmic frequency resolution. Note that the wavelet is a highpass filter that "shaves off" the upper half of the spectrum at each stage. Due to the subsampling by 2 at each stage, the time resolution decreases accordingly. That is, we have a sampling of the time/frequency plane exactly as in the continuous time wavelet transform with $a_0 = 2$. The basic building block of such a DWT is a two channel filter bank subsampled by 2. If the filter bank has the perfect reconstruction (PR) property (that is, there exists a stable PR synthesis filter bank), then the DWT can be inverted since each stage can be inverted. Furthermore, if the filter bank is lossless [4], then the filter impulse responses and their translates form an orthonormal set [7]. Since the DWT is made of a cascade of such elementary lossless banks, it is easy to verify that resulting basis is orthonormal as well.

While the theory of PR filter banks is an adequate framework for deriving DWT's implementable with filters having rational transfer functions, the wavelet theory leads to some interesting questions. One such question is the regularity of the infinitely iterated and subsampled lowpass filter appearing as the horizontal branch of the DWT in figure 1. This filter has a z-transform equal to:

$$H_\infty(z) = \prod_{i=O}^{\infty} H_0(z^{2^i}) \qquad (3)$$

I.Daubechies [2] gives a sufficient condition under which $H_\infty(z)$ converges to a continuous function (otherwise, it has fractal behavior). The condition is that the filter has a sufficient number of zeros at $z = -1$ so as to attenuate the supremum of the magnitude of the Fourier transform of the remaining factor [2]. The regularity of orthogonal wavelets with compact support (that is, lossless two channel FIR filter banks) is investigated in [2], and a construction is

CH2847-2/90/0000-1723 $1.00 © 1990 IEEE

given for shortest filters with a given regularity.

In this paper, we will look at linear phase wavelets with compact support instead. Since there are no linear phase orthogonal wavelets (except the trivial Haar case), we have to use biorthogonal bases. Now the regularity of both analysis and synthesis has to be checked, and we derive such regular pairs. Multiresolution schemes, first introduced in computer vision and then in image compression [1], are briefly revisited in the light of the DWT, showing how redundancies can be removed. Then, we discuss the complexity of the DWT and show how fast convolution techniques can decrease the number of operations. Another benefit of these FFT based techniques is that they lead naturally to a wavelet based spectrogram.

2 Biorthogonal Systems

In the two-channel critically sampled case for perfect reconstruction with FIR filters it is necessary and sufficient that :

$$\begin{pmatrix} G_0(z) \\ G_1(z) \end{pmatrix} = \frac{z^{-k_1}}{\Delta(z)} \begin{pmatrix} H_1(-z) \\ -H_0(-z) \end{pmatrix} \quad (4)$$

and $\Delta(z)$ is equal to:

$$H_0(z)H_1(-z) - H_0(-z)H_1(z) = P(z) - P(-z) = cz^{-2k_2-1} \quad (5)$$

Obviously (5) implies that while $P(z)$ can have arbitrary even coefficients it must have one and only one non-zero coefficient of an odd power of z. The $H_0(z)$ filter and subsample operation can be written as a matrix H_0 equal to:

$$\begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & h_0(L-1) & \dots & h_0(0) & 0 & 0 \\ 0 & 0 & 0 & h_0(L-1) & \dots & h_0(0) \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (6)$$

So we write : $H_0x = y$ where x and y are infinite input and output vectors respectively. Similarly the upsample and filter operation can be represented by a matrix G_0^* ($*$ stands for hermitian transpose) where a column contains $h_1(0), -h_1(1), h_1(2)$ etc, where we used used (4). H_1 and G_1^* are defined similarly. The rows of H_0 (resp. H_1) are given by the coefficients of $H_0(z)$ (resp. $H_1(z)$); and the columns of G_0^* (resp. G_1^*) are given by the coefficients of $H_1(-z)$ (resp. $-H_0(-z)$). Noting that the polynomials $H_0(z)H_0(-z)$ and $H_1(z)H_1(-z)$ each have all coefficients of odd powers of z equal to zero gives the matrix identities:

$$H_0G_1^* = 0 = H_1G_0^* \quad (7)$$

Similarly using (5) to note that $H_0(z)H_1(-z)$ has a single non-zero coefficient of an odd power of z one finds :

$$H_0G_0^* = I = H_1G_1^* \quad (8)$$

Since we have a perfect reconstruction system we get:

$$G_0^*H_0 + G_1^*H_1 = I \quad (9)$$

Note that (8) implies that $G_0^*H_0$ and $G_1^*H_1$ are each projections, onto subspaces which are not in general orthogonal. Because of (7) and (8) the analysis synthesis system is termed biorthogonal. In the special case where we have a paraunitary solution one finds: $G_0 = H_0$ and $G_1 = H_1$ and (7) gives that we have projections onto subspaces which are mutually orthogonal [3]. The above discussion shows that in a pyramid scheme, one can essentially always subsample the difference signal (ignoring quantization of course). The lowpass version corresponds to H_0. The interpolation corresponds to $G_0^*H_0$. Thus, the difference signal is $I - G_0^*H_0 = G_1^*H_1$ following (9). Thus, applying H_1 to the difference signal leads to $H_1G_1^*H_1$ which equals H_1 by (8), that is, we have now the same signals as in a subband coding system, and can therefore perfectly reconstruct. We have assumed that we can find a complementary filter H_1 to H_0 so as to achieve perfect reconstruction, and this is possible in most cases [5].

3 Regular Linear Phase Filters

For a paraunitary solution the synthesis filters are equal to the analysis filters to within a time reversal; hence both are regular provided that $H_0(z)$ is. If linear phase filters are desired we have seen that $G_0(z) = cz^{-l}H_1(-z)$ and $G_1(z) = -cz^{-l}H_0(-z)$; this implies that both $H_0(z)$ and $H_1(-z)$ must satisfy the regularity condition, since they are the iterated filters. To construct a linear phase scheme, it suffices therefore to find some $P(z) = H_0(z)H_1(-z)$, satisfying (5), where $H_0(z) = (1+z^{-1})^{N_0}R_0(z)$ and $H_1(-z) = (1+z^{-1})^{N_1}R_1(z)$, are regular, and $R_0(z)$ and $R_1(z)$ are linear phase. The idea is now to give as many zeroes at $z = -1$ to $P(z)$ as possible, leading to the most regular solution of a given degree.

For a given order $N = N_0 + N_1$ we can solve for the coefficients of $R(z) = R_0(z)R_1(z)$ by solving a set of linear equations. Assume that the degree of $R(z)$ is M; so that of $P(z)$ is $N+M$. Since $R(z)$ is linear phase it has only $\lfloor \frac{M+2}{2} \rfloor$ independent coefficients. Also note that $P(z)$ has $\lfloor \frac{N+M+1}{2} \rfloor$ odd powers of z^{-1}, and this is the number of equations to be solved. However since the binomial coefficients are symmetric, only about half of these equations are independent. In fact we have only

$$\lfloor (\lfloor (N+M+1)/2 \rfloor + 1)/2 \rfloor \quad (10)$$

independent equations to solve. Next observe that $M+N$ must be even; for if it is odd then the highest power of z^{-1} in $P(z)$ - and hence in $R(z)$ - must equal zero (it cannot be the single non-zero coefficient because of symmetry). Consider now the two cases for N:

N even: clearly M is also even giving $\lfloor \frac{M+2}{2} \rfloor = M/2+1$ coefficients. For $M = N-2$ we find from (10) that there are $N/2$ equations, as well as $N/2$ coefficients. This implies that we can find $R(z)$ by solving a system of $N/2$ equations in $N/2$ unknowns.

N odd: when N is odd so is M, which implies that $R(z)$ has a single zero at $z = 1$ or $z = -1$, since it is linear phase. However since (5) implies that $(1-z^{-1})$ cannot be a factor of $P(z)$, so the single zero must be of the form $(1+z^{-1})$.

In fact it turns out that the even case with $M = N-2$ gives the same solution as the odd case with $M' = N'$ for $N' = N-1$. That is, the effect of removing one factor $(1+z^{-1})$ from the binomial part of $P(z)$ with N even is to place that factor in $R(z)$ for $N' = N-1$ odd.

Example: if $N = 5$ and hence $M = 5$ we solve the 3 by 3 system found by imposing the constraints on the coefficients of the odd powers of z^{-1} of

$$(r_0 + r_1 z^{-1} + r_2 z^{-2} + r_2 z^{-3} + r_1 z^{-4} + r_0 z^{-5})$$
$$\cdot(1 + 5z^{-1} + 10z^{-2} + 10z^{-3} + 5z^{-4} + z^{-5}) = P(z)$$

So we solve:

$$\begin{pmatrix} 5 & 1 & 0 \\ 10 & 10 & 6 \\ 2 & 10 & 20 \end{pmatrix} \begin{pmatrix} r_0 \\ r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \quad (11)$$

Then we factor $R(z)$ into linear phase components, each of which must be made regular by adding zeros at $z = -1$ from the binomial portion of $P(z)$. Each of the filters so constructed for odd $N \leq 18$ have been found to have a regular linear phase factorization. For small N it may happen that ensuring regularity requires an inequitable distribution of the zeros. For example with $N = 5$ we find: $H_0(z) = 1 + z^{-1}$ and $H_1(-z) = (1+z^{-1})^3 R(z)$. The difficulty in finding regular factorizations appears to ease as N increases. Figure 2 shows regular analysis and synthesis wavelets for the case $N = 17$. Both filters are of length 18 and have 9 zeroes at $z = -1$ each, that is, they have the first 9 moments at the origin equal to zero.

Comment: A second approach to calculating $R(z)$ is based on the linear phase lattice form of [7]. These lattices produce linear phase perfect reconstruction filters of odd degree on both upper and lower branches. If we solve for the lattice coefficients to make the upper branch equal $(1+z^{-1})^N$ for N odd, then the lower branch must equal $R(z)$ since this is the unique linear phase polynomial of degree $M = N$ giving a $P(z)$ that satisfies (5). Note that all coefficients of $R(z)$ in both constructions are rational, making implementations simple.

4 Complexity of the DWT

The efficiency of the DWT comes from the fact that if one stage requires complexity C per input sample, then, because of the subsampling, the next stage takes $C/2$, and so on, that is, the whole DWT has complexity less than $2C$. A direct implementation with filters of length L_f leads to $C = (L_f\mu, (L_f-1)\alpha)$ per input sample (μ and α stands for multiplication and addition respectively). The operation performed by a two channel filter bank subsampled by 2 can be written in z-transform domain as:

$$\begin{pmatrix} Y_0(z) \\ Y_1(z) \end{pmatrix} = \begin{pmatrix} H_{00}(z) & H_{01}(z) \\ H_{10}(z) & H_{11}(z) \end{pmatrix} \cdot \begin{pmatrix} X_0(z) \\ X_1(z) \end{pmatrix} \quad (12)$$

where $X_j(z)$ and $H_{ij}(z)$ are the j-th polyphase component of $X(z)$ and $H_i(z)$ respectively (that is, $X(z) = X_0(z^2) + z^{-1}X_1(z^2)$, and similar relations for $H_i(z)$). When $H_i(z)$ are FIR filters in particular, the above product can be sped up by performing the polynomial products in Fourier domain. The degree of $H_{ij}(z)$ being $L_f/2-1$, take an (even) signal length L_s and a Fourier transform length L_x so that $L_x \geq L_f/2 + L_s/2 - 1$ and thus wrap-around effects are avoided. The complexity of (12) becomes equal to 2 FFT(L_x), 4 freq. domain convolutions and summation of spectras, and 2 IFFT(L_x). For the sake of discussion, take a length $L_x = 2^m$ FFT that uses typically $2^{m-1}(m-3)+2$ multiplications, freq. domain convolution using $3L_x/2 = 3\cdot 2^{m-1}$ multiplications (real input and filters) and then the complexity of (12) is $(m2^{m+1}+8)\mu$ per L_s samples, where $L_s = 2(L_x - L_f/2 + 1)$ typically (assuming an infinite input that can be segmented arbitrarily). Given a certain filter length L_f, there is an optimal L_x that will minimize multiplications per input sample. We will assume L_f to be a power of 2 and choose $L_x = L_f$ (other sizes can give an improvement, but this is sufficient for our discussion). Then, the number of multiplies per sample becomes:

$$(2L_f \log_2 L_f + 8)/(L_f + 2) \quad (13)$$

that is, we have replaced the $O[L_f]$ complexity of the direct computation by $O[\log L_f]$. For ex., when $L_f = 32$, then (13) leads to 70% savings over the direct method.

A further simplification occurs if we compute more than one stage at once in the Fourier domain. After the IFFT, one would derive even and odd indexed samples (see (12)) to enter the next stage. But if $X(k)$ is the transform domain vector of length N, then we can write:

$$x(2n) = 1/N \sum_{k=0}^{N/2-1} (X(k) + X(k+N/2)) W_{N/2}^{-kn}$$
$$x(2n+1) = 1/N \sum_{k=0}^{N/2-1} W_N^{-k}(X(k) - X(k+N/2)) W_{N/2}^{-kn}$$

where $W_N = e^{-j2\pi/N}$. That is, both $x(2n)$ and $x(2n+1)$ can be obtained from a half sized IFFT. Now, compute the next stage with an FFT of size $N/2$, so as to cancel the IFFT. The multiplication by W_N^{-k} in the expression for $x(2n+1)$ can be merged with the next convolution. Therefore, subsampling in Fourier domain is achieved at the cost of additions only. As long as the size $N/2$ transform is of sufficient size (so as to avoid wrap-around effects), we can compute the next convolution in the Fourier domain as well. For ex., compute 2 stages of a DWT with $L_f = 32$ and $L_x = 128$. Then 166 input samples still produce a valid linear convolution. The cost of the first stage is 2 FFT(128), 2 freq. domain convolutions and 1 IFFT(128). The second stage uses 2 freq. domain convolutions and 2 IFFT(64). This amounts to a total of 1542 μ's, or 9.31 μ's per input sample. A stage by stage computation would use 25 to 50% more multiplications. The savings are moderate (eventhough we avoided back and forth Fourier transforms) because the filter in the lowest stage correspond to a long filter in the input stage (due to subsampling), reducing the usable input size accordingly. Another benefit of the Fourier domain computation appears when we want to display the power spectrum instead of the time-domain waveform after filtering. Then, the IFFT's are avoided, and one obtains a wavelet based spectrogram with (block) logarithmic frequency resolution.

5 Conclusion and Directions

We have shown how to construct regular biorthogonal bases corresponding in particular to linear phase wavelets. Examples of interest have been derived. The complexity of the DWT has been investigated.

Possible directions of further research include DWT's using filter banks having rational sampling rates (leading to finer frequency resolution) as well multidimensional DWT's using non-separable decompositions and filters (allowing greater freedom than separable DWT's).

In conclusion, the crossfertilization between ideas from wavelet theory and multirate filter banks appears to be quite useful, and multiresolution signal analysis suggests numerous applications that can benefit from this unified framework.

Acknowledgements: This research was supported in part by NSF under Grants CDR-84-21402 and MIP-88-08277. Independently of this work, results on biorthogonal bases were derived by I.Daubechies, whom the authors would like to thank for helpful comments.

References

[1] P.J.Burt and E.H.Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. on Com., Vol. 31, No.4, April 1983, pp.532-540.

[2] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," Commun. on Pure and Applied Mathematics, Vol.XLI 909-996, 1988

[3] G.Strang, "Wavelets and dilation equations: a brief introduction," SIAM Review, to appear.

[4] P.P.Vaidyanathan, "Quadrature Mirror Filter Banks, M-band Extensions and Perfect-Reconstruction Technique," IEEE ASSP Magazine, Vol. 4, No. 3, pp.4-20, July 1987.

[5] M.Vetterli, "Filter Banks Allowing Perfect Reconstruction," Signal Processing, Vol.10, No.3, April 1986, pp.219-244.

[6] M.Vetterli, "A Theory of Multirate Filter Banks," IEEE Trans. on ASSP, Vol. 35, No. 3, pp. 356-372, March 1987.

[7] M.Vetterli and D. Le Gall, "Perfect reconstruction FIR filter banks: some properties and factorizations," IEEE Trans. on ASSP, Vol. 37, No. 7, July 1989, pp.1057-1071.

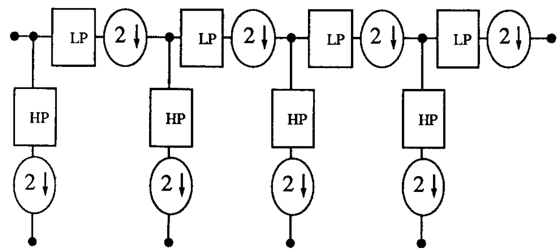

Figure 1: Discrete wavelet transform implemented with multirate filter banks.

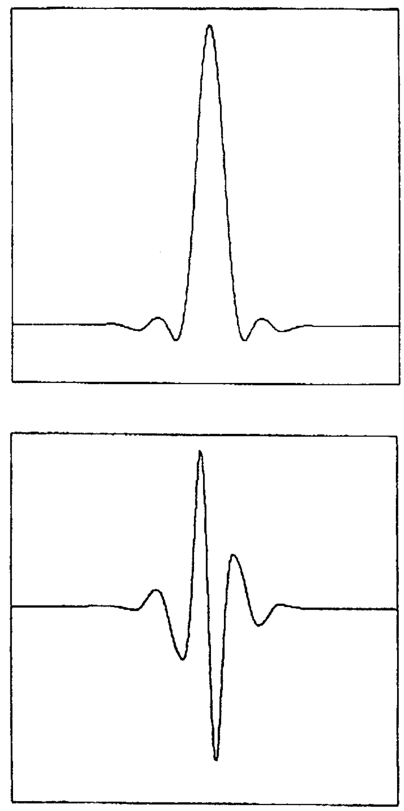

Figure 2: Regular symmetric scaling function (top) and corresponding antisymmetric wavelet (bottom) for the analysis with $N = 17$. Both are length 18 linear phase FIR filters. The equivalent scaling function and wavelet for the synthesis are linear phase and regular as well, and look similar (but are not equal to the above).